

Lecture : Local stretching theories

Jean-Luc Thiffeault

Notes by David Goluskin and Anubhab Roy

24 June 2010

1 Introduction

In this lecture we will try to understand *local* or *Lagrangian theories* involved in mixing of passive scalars. This involves solving the advection-diffusion (AD) equation along fluid trajectories. The origins of these local theories can be traced to Batchelor's idea of describing the flow via spatially-constant strain-rate matrices with prescribed time dependence [2]. Kraichnan addressed the problem next by considering the velocity field in the AD equation to be a stochastic Gaussian field with a time correlation that decays infinitely rapidly (or is white in time) and a spatial correlation that has a power-law structure [6]. This led to solving a stochastic differential equation. Zeldovich encountered the problem in the context of heat diffusion and the magnetic dynamo and adopted a random matrix theory approach [13–15]. More recently, tools from large deviation theory and path integration have aided in obtaining a complete solution of the problem, as will be discussed in this lecture.

Let us revisit the AD equation,

$$\partial_t \theta + \mathbf{u} \cdot \nabla \theta = \kappa \nabla^2 \theta, \quad \Omega = \mathbb{R}^2 \text{ or } \mathbb{R}^3. \quad (1)$$

As mentioned before $\theta(\mathbf{x}, t)$ denotes a passive scalar field (say concentration of dye) and $\mathbf{u}(\mathbf{x}, t)$ is the ambient velocity field which stirs the scalar as it diffuses with a diffusivity κ .

Previously (Lecture 1) we have examined how a “patch” of dye evolves in an extensional flow, $\mathbf{u} = (\lambda x, -\lambda y)$. We will now address the evolution of a passive scalar when acted upon by an arbitrary linear flow. By definition, for a linear flow the velocity field depends linearly on the position. Rigid body rotation, extensional flow, and simple shear are some classic examples of linear flow. In general,

$$\mathbf{u}(\mathbf{x}, t) = \mathbf{U}(t) + \mathbf{x} \cdot \mathbf{A}(t), \quad \text{Tr}(\mathbf{A}) = 0, \quad (2)$$

where $\mathbf{A}(t)$ is the velocity gradient tensor. The requirement for it to be traceless is a consequence of incompressibility of the flow field ($\nabla \cdot \mathbf{u} = 0$). For turbulent or Lagrangian-chaotic flows, $\mathbf{A}(t)$ is a random matrix, having a finite correlation time, which is the Lagrangian correlation time of the velocity. Here $\mathbf{U}(t)$ is a uniform flow. The advantage of using linear flows lies in making the problem analytically tractable and most importantly it serves as a nice prototype for more complicated flows. We will simplify things further by considering

Report Documentation Page			Form Approved OMB No. 0704-0188		
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 24 JUN 2010		2. REPORT TYPE		3. DATES COVERED 00-00-2010 to 00-00-2010	
4. TITLE AND SUBTITLE Local Stretching Theories				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Woods Hole Oceanographic Institution, Woods Hole, MA, 02543				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES See also ADA544302					
14. ABSTRACT In this lecture we will try to understand local or Lagrangian theories involved in mixing of passive scalars. This involves solving the advection-diffusion (AD) equation along uid trajectories. The origins of these local theories can be traced to Batchelor's idea of describing the flow via spatially-constant strain-rate matrices with prescribed time dependence [2]. Kraichnan addressed the problem next by considering the velocity field in the AD equation to be a stochastic Gaussian field with a time correlation that decays in time rapidly (or is white in time) and a spatial correlation that has a power-law structure [6]. This led to solving a stochastic differential equation. Zeldovich encountered the problem in the context of heat diffusion and the magnetic dynamo and adopted a random matrix theory approach [13][15]. More recently, tools from large deviation theory and path integration have aided in obtaining a complete solution of the problem, as will be discussed in this lecture.					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 12	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

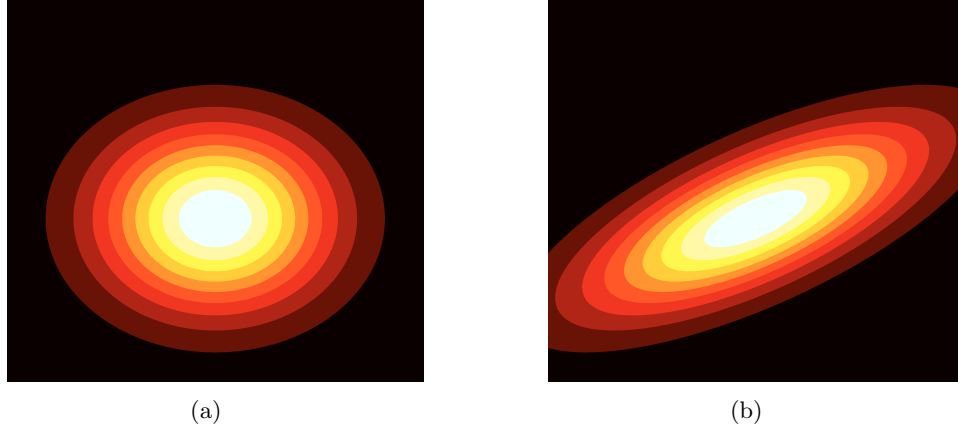


Figure 1: Advection of a concentration field by a simple shear flow, $\mathbf{u} = (U_x, U_y) = (\dot{\gamma}y, 0)$ (neglecting diffusion, $\dot{\gamma}$ assumed constant). An initially isotropic concentration field (a) gets both stretched and rotated by the background flow (b).

the limit of high Schmidt number, which is a dimensionless quantity defined as

$$Sc := \nu/\kappa$$

where ν is the kinematic viscosity of the fluid and κ as already mentioned is the diffusivity of the scalar. (If the scalar is heat, then the Schmidt number becomes the Prandtl number.) Momentum diffuses significantly faster than the passive scalar for flows with large Sc . This is the desired limit for chaotic advection: $Sc \gg 1$ ensures that the velocity field appears much smoother than the passive scalar concentration, allowing the linear local approximation (2). This is the regime that was studied by Batchelor and leads to the celebrated Batchelor spectrum. For $Sc \leq 1$ we have the Kolmogorov-Obukhov-Corrsin regime, where the scalar gets advected by the turbulent inertial range, unlike the Batchelor regime. (We will not be discussing the KOC regime in this lecture.)

2 ‘Inertia tensor’ of a patch of passive scalar

It is of great interest to understand the response of a passive scalar to a flow with chaotic Lagrangian trajectories. As already discussed, a decent approximation to such a velocity field can be made by superposing random linear flows (in the limit $Sc \gg 1$). In order to get a statistical description of the passive scalar field subjected to such an ensemble of linear flows, we need to understand the behaviour of the moments of $\theta(\mathbf{x}, t)$. We use angle brackets to denote the volume-integrated value of a quantity (similar to an average),

$$\langle f \rangle = \int_{\Omega} f dV.$$

We have already seen (Lecture 1) that when the domain is either periodic or is bounded by insulating walls then

$$\partial_t \langle \theta \rangle = 0, \tag{3}$$

that is, the mean value of θ is conserved. This is a consequence of the no-net-flux boundary condition imposed on the system. To track the evolution of ‘a blob of dye,’ one can define the ‘center of mass’ of θ as

$$c_i = \frac{\langle x_i \theta \rangle}{\langle \theta \rangle}. \quad (4)$$

To obtain an evolution equation for \mathbf{c} we consider the first moment of the AD equation, and integrate by parts a few times:

$$\begin{aligned} \partial_t \langle x_i \theta \rangle + \langle x_i \nabla \cdot ((\mathbf{U} + \mathbf{x} \cdot \mathbf{A}) \theta) \rangle &= \kappa \langle x_i \nabla^2 \theta \rangle, \\ \Rightarrow \partial_t \langle x_i \theta \rangle - \langle (U_j + x_\ell A_{\ell j}) \theta \partial_j x_i \rangle &= 0, \\ \Rightarrow \partial_t \langle x_i \theta \rangle - U_i \langle \theta \rangle - A_{\ell i} \langle \theta \rangle c_\ell &= 0, \quad (\partial_j x_i = \delta_{ij}) \end{aligned}$$

which leads to

$$\dot{\mathbf{c}} = \mathbf{U} + \mathbf{c} \cdot \mathbf{A}. \quad (5)$$

Here the overdot represents a derivative with respect to time. In the calculation the contribution due to the dissipative term disappeared due to the no-net-flux boundary conditions. Thus the center of mass of the patch gets advected by the background flow.

Proceeding to the calculation for the next moment we define the ‘inertia tensor’ of the patch as

$$M_{ij} = \frac{\langle x_i x_j \theta \rangle}{\langle \theta \rangle} - c_i c_j.$$

The additional $-c_i c_j$ term makes \mathbf{M} translationally invariant. By construction \mathbf{M} is also a symmetric matrix. We proceed in similar fashion as above to obtain the corresponding evolution equation for \mathbf{M} :

$$\begin{aligned} \partial_t \langle x_i x_j \theta \rangle + \langle x_i x_j \nabla \cdot ((\mathbf{U} + \mathbf{x} \cdot \mathbf{A}) \theta) \rangle &= \kappa \langle x_i x_j \nabla^2 \theta \rangle \\ \Rightarrow \partial_t \langle x_i x_j \theta \rangle - U_j \langle x_i \theta \rangle - U_i \langle x_j \theta \rangle - A_{pj} \langle x_i x_p \theta \rangle - A_{pi} \langle x_j x_p \theta \rangle &= 2\kappa \delta_{ij} \langle \theta \rangle \\ \Rightarrow \partial_t (M_{ij} + c_i c_j) - U_j c_i - U_i c_j - A_{pj} (M_{ip} + c_i c_p) - A_{pi} (M_{jp} + c_j c_p) &= 2\kappa \delta_{ij} \\ \Rightarrow \partial_t M_{ij} + c_i (\partial_t c_j - U_j - c_p A_{pj}) + c_j (\partial_t c_i - U_i - c_p A_{pi}) &= M_{ip} A_{pj} + A_{pi} M_{jp} + 2\kappa \delta_{ij} \end{aligned}$$

so that finally

$$\dot{\mathbf{M}} = \mathbf{M} \cdot \mathbf{A} + \mathbf{A}^\dagger \cdot \mathbf{M} + 2\kappa \mathbf{I}. \quad (6)$$

Thus we have obtained an evolution equation for the ‘inertia tensor’ of the patch. Here \mathbf{A}^\dagger denotes the transpose of the velocity gradient tensor.

The solution for $\mathbf{c}(t)$ and $\mathbf{M}(t)$ can be explicitly written as

$$\mathbf{c}(t) = \mathbf{c}(0) \cdot e^{\int_0^t \mathbf{A}(\tau_1) d\tau_1} + \int_0^t \mathbf{U}(\tau) \cdot e^{\int_\tau^t \mathbf{A}(\tau_1) d\tau_1} d\tau, \quad (7)$$

$$\mathbf{M}(t) = e^{\int_0^t \mathbf{A}^\dagger(\tau_1) d\tau_1} \cdot \mathbf{M}(0) \cdot e^{\int_0^t \mathbf{A}(\tau_1) d\tau_1} + 2\kappa \int_0^t e^{\int_\tau^t \mathbf{A}^\dagger(\tau_1) d\tau_1} \cdot e^{\int_\tau^t \mathbf{A}(\tau_1) d\tau_1} d\tau. \quad (8)$$

Here the quantity $e^{\mathbf{B}}$ is a matrix exponential, defined by the usual power series

$$e^{\mathbf{B}} = \mathbf{I} + \mathbf{B} + \frac{1}{2} \mathbf{B} \cdot \mathbf{B} + \dots$$

It should be noted that in the expression (8) for $\mathbf{M}(t)$ the two matrix exponentials cannot be combined and written as $e^{\int_{\tau}^t (\mathbf{A}^\dagger + \mathbf{A})(\tau_1) d\tau_1}$. For matrix exponentials, $e^{\mathbf{B}_1 + \mathbf{B}_2} = e^{\mathbf{B}_1} \cdot e^{\mathbf{B}_2}$ only if the matrices \mathbf{B}_1 and \mathbf{B}_2 commute. The velocity gradient matrix, \mathbf{A} , is typically a non-normal matrix ($\mathbf{A}\mathbf{A}^\dagger \neq \mathbf{A}^\dagger\mathbf{A}$).

For a simple shear flow, we have $\mathbf{u} = (\dot{\gamma}y, 0)$ and

$$\mathbf{A} = \begin{pmatrix} 0 & 0 \\ \dot{\gamma} & 0 \end{pmatrix}.$$

Since \mathbf{A} is nilpotent ($\mathbf{A}^2 = 0$), the matrix exponential has only two terms in its power series representation, so that

$$e^{\mathbf{A}t} = \mathbf{I} + \mathbf{A}t = \begin{pmatrix} 1 & 0 \\ \dot{\gamma}t & 1 \end{pmatrix}.$$

3 Computation of Lyapunov exponents

Since \mathbf{M} is a symmetric matrix the eigendecomposition takes the simple form

$$\mathbf{M} = \mathbf{R}\mathbf{D}\mathbf{R}^\dagger \quad (9)$$

where \mathbf{R} is an orthogonal matrix (the columns of which are eigenvectors of \mathbf{M}), and \mathbf{D} is real and diagonal (having the eigenvalues of \mathbf{M} on the diagonal). When the eigenvalues are sufficiently distinct, the radial and angular degrees of freedom decouple from each other, as we will now show. (Our derivation parallels that of [1].) Let us substitute the expression (9) in the evolution equation (8):

$$\begin{aligned} \dot{\mathbf{R}}\mathbf{D}\mathbf{R}^\dagger + \mathbf{R}\mathbf{D}\dot{\mathbf{R}}^\dagger + \mathbf{R}\dot{\mathbf{D}}\mathbf{R}^\dagger &= \mathbf{R}\mathbf{D}\mathbf{R}^\dagger\mathbf{A} + \mathbf{A}^\dagger\mathbf{R}\mathbf{D}\mathbf{R}^\dagger + 2\kappa\mathbf{I}, \\ \mathbf{R}^\dagger\dot{\mathbf{R}}\mathbf{D} + \mathbf{D}\dot{\mathbf{R}}^\dagger\mathbf{R} + \dot{\mathbf{D}} &= \mathbf{D}\mathbf{R}^\dagger\mathbf{A}\mathbf{R} + \mathbf{R}^\dagger\mathbf{A}^\dagger\mathbf{R}\mathbf{D} + 2\kappa\mathbf{I}. \end{aligned} \quad (10)$$

Since

$$\frac{d}{dt}(\mathbf{R}^\dagger\mathbf{R}) = \dot{\mathbf{R}}^\dagger\mathbf{R} + \mathbf{R}^\dagger\dot{\mathbf{R}} = 0$$

both $\dot{\mathbf{R}}^\dagger\mathbf{R}$ and $\mathbf{R}^\dagger\dot{\mathbf{R}}$ are antisymmetric matrices. Thus in equation (10) the first two terms on the left do not have any diagonal elements. Denoting $\tilde{\mathbf{A}} = \mathbf{R}^\dagger\mathbf{A}\mathbf{R}$ as the rotated velocity gradient matrix we have the following equation for the diagonal elements,

$$\dot{D}_{ii} = 2\tilde{A}_{ii}D_{ii} + 2\kappa,$$

where repeated indices do not indicate summation. If we assume $D_{ii} = e^{2\rho_i}$ ($\rho_i > 0$ denotes extension whereas $\rho_i < 0$ represents compression) then one has

$$\dot{\rho}_i = \tilde{A}_{ii} + \kappa e^{-2\rho_i}. \quad (11)$$

Without loss of generality, we arrange the eigenvalues $e^{2\rho_1}, e^{2\rho_2}, \dots, e^{2\rho_d}$ in non-increasing order:

$$\rho_1 \geq \rho_2 \cdots \geq \rho_d \quad (12)$$

where d is the dimension of space.

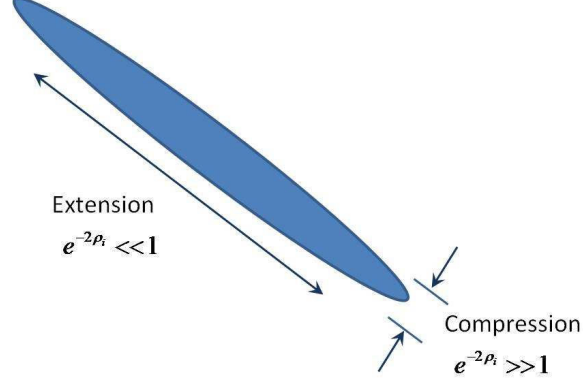


Figure 2: The diffusive term ($\kappa e^{-2\rho_i}$) in the equation for ρ_i would be negligible along the extensional axis compared to the compressional direction.

Having worked with the diagonal matrix, \mathbf{D} , let us focus on the eigenvector matrix, \mathbf{R} . Considering the off-diagonal terms of equation (10),

$$\begin{aligned} [\mathbf{R}^\dagger \dot{\mathbf{R}} \mathbf{D}]_{ij} &= [\mathbf{R}^\dagger \dot{\mathbf{R}}]_{i\ell} D_{\ell j} = [\mathbf{R}^\dagger \dot{\mathbf{R}}]_{ij} D_{jj} \\ [\mathbf{D} \dot{\mathbf{R}}^\dagger \mathbf{R}]_{ij} &= D_{i\ell} [\dot{\mathbf{R}}^\dagger \mathbf{R}]_{\ell j} = -[\mathbf{R}^\dagger \dot{\mathbf{R}}]_{ij} D_{ii} \end{aligned}$$

where repeated indices do not indicate summation. Let us define $\mathbf{R}^\dagger \dot{\mathbf{R}} = \mathbf{\Omega}$, an antisymmetric matrix. Thus we have,

$$\begin{aligned} (D_{jj} - D_{ii})\Omega_{ij} &= D_{ii}\tilde{A}_{ij} + \tilde{A}_{ji}D_{jj} \\ \Omega_{ij} &= \frac{e^{2\rho_i}\tilde{A}_{ij} + e^{2\rho_j}\tilde{A}_{ji}}{e^{2\rho_j} - e^{2\rho_i}}, \quad i \neq j \end{aligned} \quad (13)$$

Let us assume there exists scale separation in ρ 's ($\rho_1 \gg \rho_2 \gg \dots \gg \rho_d$).¹ Thus if $\rho_i \gg \rho_j$, the $t \gg 1$ solution is

$$\Omega_{ij} \approx \begin{cases} -\tilde{A}_{ij}, & i < j \\ \tilde{A}_{ji}, & i > j \end{cases} \quad (14)$$

Thus $\tilde{\mathbf{A}}$ is independent of the eigenvalues and we can solve equation (11) directly,

$$\rho_i(t) = \rho_i(0) + \mathcal{A}_i(t) + \frac{1}{2} \log \left[1 + 2\kappa e^{-2\rho_i(0)} \int_0^t e^{-2\mathcal{A}_i(t')} dt' \right] \quad (15)$$

where $\mathcal{A}_i(t) = \int_0^t \tilde{A}_{ii}(\tau) d\tau$.

When diffusion vanishes ($\kappa = 0$), this reduces to

$$\rho_i(t) = \rho_i(0) + \int_0^t \tilde{A}_{ii}(\tau) d\tau. \quad (16)$$

¹This is almost always true for long times, but it may fail locally or due to a symmetry in the problem.

The $\kappa = 0$ case can be used to compute Lyapunov exponents² from the eigenvalues of \mathbf{M} :

$$\lambda_i = \lim_{t \rightarrow \infty} \frac{\log(e^{\rho_i(t)})}{t} = \lim_{t \rightarrow \infty} \frac{\rho_i(t)}{t} \quad (17)$$

The convergence of the above limit is assured by Oseledec's multiplicative ergodic theorem [7].

A blob of initial size L will become an ellipsoid with the length of its main axis changing as $\exp(\lambda_i t)$ along directions corresponding to positive and negative Lyapunov exponents. The rate of decrease of the smallest dimension is determined by λ_d . Because of diffusion, the contraction terminates after a time, $t_\kappa \approx |\lambda_d|^{-1} \log(L^2 |\lambda_d| / \kappa)$, when the diffusive lengthscale, $r_\kappa = \sqrt{\kappa / |\lambda_d|}$ is attained.

4 Large-deviation form of the central limit theorem³

To form a probabilistic theory of mixing, we make \mathbf{A} , and therefore ρ_1 and ρ_2 (we restrict to $d = 2$ dimensions from now), into random variables and deduce the approximate long-time moments of concentration. To do so, we require a central limit theorem (CLT) which is also valid for large deviations.

Let x_i be i.i.d. (independent and identically distributed) random variables with finite mean \bar{x} and variance σ^2 . Let x be the mean of n such x_i . Roughly speaking, the classical CLT asserts that

$$P(x, n) \sim \frac{1}{\sqrt{2\pi n \sigma^2}} \exp\left(-\frac{n(x - \bar{x})^2}{2\sigma^2}\right) \text{ when } n \gg 1 \text{ and } x - \bar{x} < \frac{\sigma}{\sqrt{n}}.$$

The theorem applies to large n , but the condition that $x - \bar{x} < \frac{\sigma}{\sqrt{n}}$ means that as n gets large, the region around the mean to which the theorem applies grows increasingly narrow. If one is concerned about the tails of the distribution, the theorem is not powerful enough. This is not surprising since it contains only second-moment information about the x_i variables, and higher moments may be critical at the tails. In the theory of mixing, the tails of the ρ PDF's are indeed significant, so we must use the large deviation form of the CLT, which is stronger and uses information from all moments of the x_i variables. The significance of the tails will be evident *a posteriori* from the analysis using the large deviation CLT. The large deviation CLT asserts that

$$P(x, n) \sim \exp(-nS(x - \bar{x})),$$

where S is the Cramér function (or rate function, or entropy function) determined by the PDF $p(x_i)$. (In these notes we typically neglect algebraic prefactors in asymptotic relations, such as the one above.) Whatever the form of $p(x_i)$, S is convex, and $S(0)$ and $S'(0)$ vanish. Taylor expanding S about 0 in the large deviation CLT, the first nonzero term is quadratic

²Lyapunov exponents describe the rate at which infinitesimally close trajectories in a dynamical system diverge from each other. A chaotic system is usually indicated by a positive infinite-time Lyapunov exponent.

³The 'facts' invoked in this section come from theorems which for the most part give bounds, not equalities. We are leaving out many details that are necessary to make the theory rigorous. See for example [3, 8, 11] for more details.

in $x - \bar{x}$. Truncating after this term recovers the classical form of the CLT, so the classical form simply approximates the large deviation form to the first nontrivial order. The explicit expression for the Cramér function in terms of $p(x_i)$ involves forward and inverse Fourier transforms, which generally cannot be explicitly evaluated. Rather than prove the large-deviation CLT or derive the general expression for the Cramér function (see e.g. [3,8,10,11]), we shall merely derive the Cramér function explicitly in the simple case where the x_i have a Bernoulli distribution, though the general formula for S may be derived by an analogous procedure.

5 Large deviation CLT for the Bernoulli distribution

Let x_i have the PDF of a fair coin,

$$p(x_i) = \frac{1}{2}\delta(x_i + 1) + \frac{1}{2}\delta(x_i - 1).$$

The characteristic function of a random variable is the Fourier transform of its PDF, generally written in the form $e^{-s(k)}$. We wish to derive $p(x, n)$, the PDF of the mean x of n random variables x_i . We are given $p(x_i)$, not $p(x, n)$, but we can express the latter in terms of the former. Because the x_i are independent, the joint distribution of x_1, \dots, x_n is simply the product of their individual PDF's, so

$$\begin{aligned} p(x, n) &= \int p(x_1, \dots, x_n) \delta\left(x - \frac{1}{n} \sum_{i=1}^n x_i\right) dx_1 \cdots dx_n \\ &= \int p(x_1) \cdots p(x_n) \delta\left(x - \frac{1}{n} \sum_{i=1}^n x_i - x\right) dx_1 \cdots dx_n. \end{aligned}$$

It is hard to compute the PDF directly from the PDF of x_i by the above formula. Rather, we shall compute the generating function of x , and then its PDF. This is easier because integrating out x -dependence allows us to then integrate out the x_i variables. The generating function of x is

$$\begin{aligned} e^{-S(k)} &= \int p(x, n) e^{-ikx} dx \\ &= \int p(x_1) \cdots p(x_n) \delta\left(x - \frac{1}{n} \sum_{i=1}^n x_i\right) e^{-ikx} dx_1 \cdots dx_n dx \\ &= \int p(x_1) \cdots p(x_n) \exp\left(-\frac{ik}{n} \sum_{i=1}^n x_i\right) dx_1 \cdots dx_n \\ &= \prod_{i=1}^n \int p(x_i) e^{-\frac{ikx_i}{n}} dx_i \\ &= \left(\int p(x_1) e^{-\frac{ikx_1}{n}} dx_1 \right)^n \\ &= \left(\frac{1}{2} \int (\delta(x_1 + 1) + \delta(x_1 - 1)) e^{-\frac{ikx_1}{n}} dx_1 \right)^n \\ &= \cos^n\left(\frac{k}{n}\right). \end{aligned}$$

We recover the PDF of x by taking the inverse Fourier transform of its generating function:

$$\begin{aligned} p(x, n) &= \frac{1}{2\pi} \int \cos^n\left(\frac{k}{n}\right) e^{ikx} dk \\ &= \frac{1}{2\pi} \int \exp\left(n \ln\left[\cos\left(\frac{k}{n}\right)\right] + ikx\right) dk \\ &= \frac{n}{2\pi} \int \exp\left(n[\ln(\cos K) + iKx]\right) dK, \end{aligned}$$

where $K \equiv \frac{k}{n}$. The last form of the above integral suggests the stationary phase approximation, since the frequency of the oscillatory part of the integrand goes to infinity as n does. Omitting the details, the stationary phase approximation yields

$$S(x) \sim -\frac{1}{2}(x+1) \ln\left(\frac{1-x}{x+1}\right) + \ln(1-x).$$

The classical form the CLT would give only

$$\begin{aligned} p(x, n) &\sim \exp\left(-\frac{N}{2} S''(0)(x - \bar{x})^2\right) \\ &\sim \exp\left(-\frac{(x - \bar{x})^2}{2/N}\right). \end{aligned}$$

In this example, the classical CLT overestimates the weight of the tails of $p(x, n)$, though typically the opposite is the case. In any case, the large deviation form is needed for accuracy at the tails.

6 Large deviation CLT applied to mixing

Recall that ρ_2 has generally become quite small at long times. However, in a fluid with diffusion, the ρ_2 dimension eventually becomes small enough that further contraction by stretching is balanced by diffusion. This is the Batchelor width, $\sqrt{\kappa/\lambda_1}$, beyond which the scalar blob will contract no further. As such late times we can assume that the PDF of ρ_2 is stationary in time, so as $t \rightarrow \infty$,

$$P(\rho_1, \rho_2, t) \sim \exp\left(-tS\left(\frac{\rho_1}{t} - \lambda_1\right)\right) P_{\text{stat}}(\rho_2).$$

The exact stationary distribution of ρ_2 is irrelevant to the heuristics of this theory. Suppose that the concentration at each point in a blob of scalar is inversely proportional to the blob's total volume, so

$$\theta(\mathbf{x}, t) \sim (\det M)^{1/2} = e^{-(\rho_1 + \rho_2)}.$$

This is an accurate assumption for a Gaussian blob, for instance. We now combine the above approximations to estimate the expectation of θ^α .

$$\begin{aligned} \langle \theta^\alpha \rangle(t) &\sim C(t) \int e^{-\alpha(\rho_1 + \rho_2)} \exp\left(-tS\left(\frac{\rho_1}{t} - \lambda_1\right)\right) P_{\text{stat}}(\rho_2) d\rho_1 d\rho_2 \\ &\sim \int e^{-\alpha\rho_1} \exp\left(-tS\left(\frac{\rho_1}{t} - \lambda_1\right)\right) d\rho_1, \end{aligned}$$

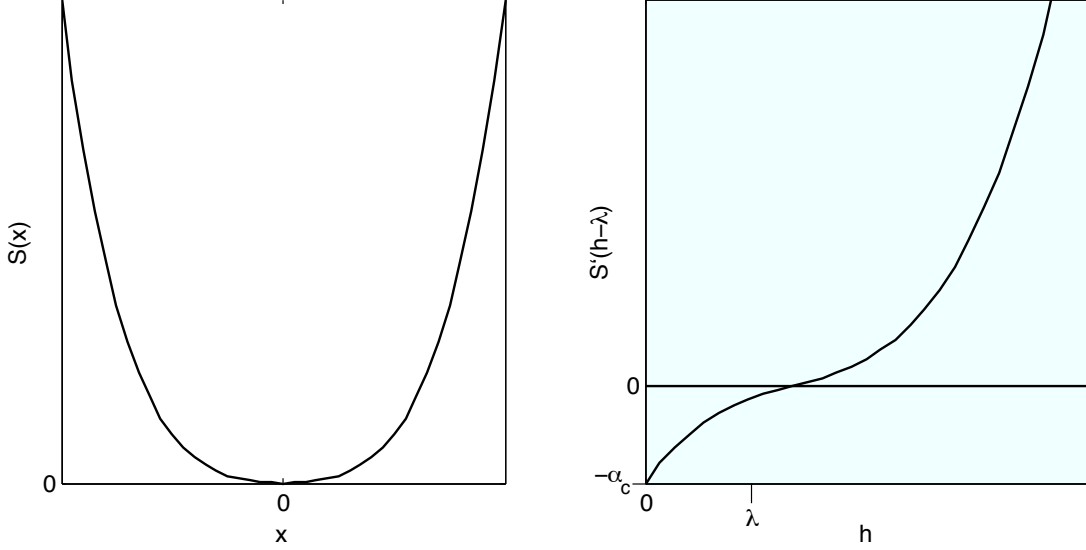


Figure 3: Generic $S(x)$ and $S(h - \lambda_1)$. When $\alpha > \alpha_c$, there is no positive h^* such that $S'(h^* - \lambda) = -\alpha$, so we must choose $h^* = 0$ for the saddle-point approximation.

where $C(t)$ is a normalization that is algebraic in t and is therefore asymptotically dominated by the integral, so we have ignored it along with whatever constant emerges from integrating over ρ_2 . Letting $h \equiv \frac{\rho_1}{t}$,

$$\langle \theta^\alpha \rangle(t) \sim \int e^{-t(\alpha h + S(h - \lambda_1))} dh = \int e^{-tH(h)} dh,$$

where $H(h) \equiv \alpha h + S(h - \lambda)$. This integral may be approximated by the saddle-point method for large t .

The saddle-point method typically asserts that

$$\langle \theta^\alpha \rangle(t) \sim e^{-H(h^*)t},$$

where h^* is such that $H'(h^*) = 0$, i.e. such that $S'(h^* - \lambda_1) = -\alpha$. In the present case, it is not always possible to choose such an h^* because h must be nonnegative. This is clear from the generic S and $S'(h - \lambda)$ curves plotted in Figure 6. Whatever the exact form of the Cramér function, $S'(h)$ begins at the origin and increases monotonically in h , so $S'(h - \lambda)$ begins at some negative point $(0, -\alpha_c)$ and increases monotonically in h . When $\alpha < \alpha_c$, $H'(h^*) = 0$ is possible, but when $\alpha > \alpha_c$, the best we can do is to choose $h^* = 0$. That is,

$$\langle \theta^\alpha \rangle(t) \sim e^{-t(\alpha h^* + S'(h^* - \lambda_1))},$$

where $h^* = \max\{0, \lambda + (S')^{-1}(-\alpha)\}$. Neglecting large deviations, the classical CLT would have given a decay rate of $-\alpha\lambda$, which is an upper bound on the decay rate we derived. In fact, the tails of the PDF of ρ_1 act significantly to slow the decay of concentration. Above α_c , $\langle \theta^\alpha \rangle(t)$ is constant in α . We may understand this plateauing as high concentration blobs dominating the expectation for large α . These are the blobs that have experienced no

stretching, so $\alpha > \alpha_c$ moments are determined solely by realizations of θ without stretching. Note that the plateauing of moments also means that the L^α norm of θ vanishes as $\alpha \rightarrow \infty$, which is characteristic of intermittency.

References

- [1] E. BALKOVSKY AND A. FOUXON, *Universal long-time properties of Lagrangian statistics in the Batchelor regime and their application to the passive scalar problem*, Phys. Rev. E, 60 (1999), pp. 4164–4174.
- [2] G. K. BATCHELOR, *Small-scale variation of convected quantities like temperature in turbulent fluid: Part 1. General discussion and the case of small conductivity*, J. Fluid Mech., 5 (1959), pp. 113–133.
- [3] R. S. ELLIS, *Entropy, Large Deviations, and Statistical Mechanics*, Springer-Verlag, New York, 1985.
- [4] G. FALKOVICH, K. GAWĘDZKI, AND M. VERGASSOLA, *Particles and fields in turbulence*, Rev. Mod. Phys., 73 (2001), pp. 913–975.
- [5] P. H. HAYNES AND J. VANNESTE, *What controls the decay of passive scalars in smooth flows?*, Phys. Fluids, 17 (2005), p. 097103.
- [6] R. H. KRAICHNAN, *Small-scale structure of a scalar field convected by turbulence*, Phys. Fluids, 11 (1968), pp. 945–953.
- [7] V. I. OSELEDEC, *A multiplicative theorem: Lyapunov characteristic numbers for dynamical systems*, Trans. Moscow Math. Soc., 19 (1968), pp. 197–231.
- [8] A. SCHWARTZ AND A. WEISS, *Large Deviations for Performance Analysis*, Chapman & Hall, London, 1995.
- [9] J. SUKHATME AND R. T. PIERREHUMBERT, *Decay of passive scalars under the action of single scale smooth velocity fields in bounded two-dimensional domains: From non-self-similar probability distribution functions to self-similar eigenmodes*, Phys. Rev. E, 66 (2002), p. 056032.
- [10] J.-L. THIFFEAULT, *Scalar decay in chaotic mixing*, in Transport and Mixing in Geophysical Flows, J. B. Weiss and A. Provenzale, eds., vol. 744 of Lecture Notes in Physics, Berlin, 2008, Springer, pp. 3–35.
- [11] S. R. S. VARADHAN, *Large deviations and applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1984.
- [12] W. R. YOUNG, *Stirring and mixing*, in Proceedings of the 1999 Summer Program in Geophysical Fluid Dynamics, J.-L. Thiffeault and C. Pasquero, eds., Woods Hole, MA, 1999, Woods Hole Oceanographic Institution. <http://gfd.whoi.edu/proceedings/1999/PDFvol1999.html>.

- [13] Y. B. ZELDOVICH, *The asymptotic law of heat transfer at small velocities in the finite domain problem*, Zhurnal eksperimentalnoi i teoreticheskoi fiziki, 7 (1937), pp. 1466–1468.
- [14] ———, *Exact solution to the problem of diffusion in a periodic velocity field and turbulent diffusion*, Doklady Akademii Nauk SSSR, 266 (1982), pp. 821–826.
- [15] Y. B. ZELDOVICH, A. A. RUZMAIKIN, S. A. MOLCHANOV, AND D. D. SOKOLOFF, *Kinematic dynamo problem in a linear velocity field*, J. Fluid Mech., 144 (1984), pp. 1–11.